ABSTRACT OF THE DISCLOSURE

In a voice browser apparatus, an HTML document is obtained from a designated URL, the obtained document is analyzed based on a rule designated from a plurality

5   of rules for defining voice input/output contents to fetch voice output contents, voice input candidates and designation information for designating a next object of processing corresponding to each voice input candidate, and the fetched voice output contents are

10   voice-outputted. Then, the voice input from a user for this output is voice-recognized, and controls obtainment of a new document and next analysis based on designation information corresponding to the input candidate matching the recognition result. Thereby,

15   contents for the graphical user interface can easily be browsed using voice input/output.